

ИВАНОВ А. Д.

КУРС НА РОБОТИЗИРОВАННУЮ ЖУРНАЛИСТИКУ: ПОЧЕМУ РОССИЙСКИМ МЕДИА НЕ ГРОЗЯТ СОВРЕМЕННЫЕ МИРОВЫЕ ТЕНДЕНЦИИ

Аннотация: Автор формулирует и анализирует технологические и экономические проблемы, препятствующие развитию роботизированной журналистики в России, в контексте появления первых роботов-журналистов в крупных международных СМИ.

Ключевые слова: роботизированная журналистика, вычислительная журналистика, автоматизированная журналистика, машинное обучение, роботы-журналисты

Ivanov A. D.

COURSE ON ROBOTIC JOURNALISM: WHY THE RUSSIAN MEDIA ARE BEHIND WORLD TENDENCIES

Abstract: The author formulates and analyzes the technological and economic problems interfering the development of robotic journalism in Russia in the context of emergence of the first robots journalists in large international media.

Keywords: robotic journalism, automated journalism, computational journalism, robotic journalist, machine learning

В очередной раз отечественная журналистика остается вне мейнстримных течений, связанных с внедрением технологий машинного обучения в рутинные процессы по сбору, обработке и публикации новостной информации в сфере массовых коммуникаций. Ведущие международные издания Англии («The Guardian») и США («Forbes», «Los Angeles Times», «Associated Press») во всю тестируют в полевых условиях роботов-журналистов, которые готовят публикации на тему происшествий, финансов и спорта [3, с. 107]. В январе 2017 года дебютную новость об ажиотаже в преддверии Чуньцзе – китайского Нового года – опубликовал медиа-робот Сяо Нань, разработанный в стенах Пекинского университета. Публикация была размещена в местной гуанчжоуской газете «Southern Metropolis Daily» и представляла собой заметку из 271 иероглифа [5]. Русскоязычная версия итоговой заметки, подготовленной и опубликованной роботом, доступна на информационно-сервисном портале «Индикатор», посвященном науке [6].

На весь процесс подготовки новости у роботов уходят считанные секунды: они моментально анализируют отчеты, оперативно собирают данные в интернете, интерпретируют их и формируют качественный журналистский текст, пригодный для размещения в СМИ. Причем тесты, проведенные разработчиками алгоритма «Narrative Science», активно использующегося в «Forbes», продемонстрировали превосходство новых технологий перед «устаревшим» человеческим трудом. Читатели не смогли распознать: находится перед ними журналистский текст или материал, подготовленный роботом [2, с. 34].

Похвастаться импортозамещением в области роботизированной журналистики пока не выходит – российские медиа не спешат тратить бюджеты на перспективные разработки ближайшего будущего и пытаются справиться с более

реальными проблемами экономического характера, а также со сменой форматов медиапотребления аудитории, которая все больше уходит в мобильный интернет. Единственными примерами внедрения так называемых нейронных сетей стали чат-боты в российской социальной сети «ВКонтакте» [4, с. 128], способные распознавать запросы пользователей и выдавать ответы, а также ML-проекты «Яндекса». Например, в 2015 году был запущен полноценный алгоритм «Яндекс для медиа», который занимается генерацией коротких журналистских текстов о загруженности дорог, состоянии погоды на основе одноименных сервисов компании – «Яндекс.Пробки» и «Яндекс.Погода». Кроме того, ТОП-5 новостей в системе «Яндекс.Новости» формируется при помощи нейросети – на основе самых популярных и актуальных публикаций отечественных СМИ алгоритм пишет уникальные заголовки новостного ТОПа. Однако о полномасштабных медийных проектах по созданию искусственного интеллекта в сфере массовых коммуникаций речи пока не идет.

На наш взгляд, сложностью в разработке искусственного интеллекта для задач медиасферы являются следующие аспекты:

- отсутствие интереса со стороны СМИ к данным разработкам вызвано в первую очередь экономическими сложностями, периодически наступающими отечественную журналистику [1, с. 9]. Разработка, обучение и запуск нейросети для задач масс-медиа – услуга очень дорогая и требующая больших временных затрат со стороны медиа;
- отсутствие так называемых размеченных корпусов журналистских текстов на русском языке, использующихся для обучения нейронных сетей, – массивов текста, разделенных на отдельные смысловые и структурные элементы с выделенными контекстными связями. Сложность обработки естественного русского языка в значительной степени тормозит развитие всей отрасли машинного обучения в России;
- большую роль в России играет консервативность подходов в журналистской деятельности – наблюдается определенная приверженность редакционных коллективов правилам и установкам классической журналистики и особенно в региональных СМИ;
- ограниченность существующих разработок с использованием нейросетей для текстогенерации. Роботы не умеют брать интервью, писать очерки и проводить глубокую аналитику с использованием эмпирического опыта – пока это очень далекие перспективы развития отрасли.

Таким образом, отечественная журналистика испытывает и технологические, и экономические сложности, которые лишают российских медиаспециалистов возможности использовать новейшие технологии машинного обучения в своей работе.

ЛИТЕРАТУРА

1. Вартанова Е. Л. О трех векторах будущей журналистики // Медиаальманах. 2015. № 3. С. 8–10.
2. Иванов А. Д. Роботизированная журналистика и первые алгоритмы на службе редакций международных СМИ // Знак: проблемное поле медиаобразования. 2016. № 2 (16). С. 32–40.

3. Иванов А. Д. Современное состояние роботизированной журналистики новостей // Журналистика цифровой эпохи: как меняется профессия. Материалы международной научно-практической конференции. 2016. С. 106–109.
4. Иванов А. Д. Чат-бот в Telegram и ВКонтакте как новый канал распространения новостей // Вестн. Волжского ун-та им. В. Н. Татищева. 2016. № 2 (16). С. 126–132.
5. Chinese robot “reporter” makes its publishing debut [Электронный ресурс]. URL: www.ecns.cn/2017/01-19/242444.shtml (дата обращения 11.03.17).
6. Indicator.Ru перевел первую статью китайского робота-журналиста [Электронный ресурс]. URL: www.indicator.ru/news/2017/01/19/vyshla-pervaya-statya-kitajskogo-robotazhurnalista/ (дата обращения 19.02.17)